

# Demystifying double-*is*

Teresa O'Neill\*

## 1 Introduction

The English double-*is* construction exemplified in (1) poses a challenge for syntactic analysis: in double-*is* sentences, it appears that two adjacent finite verbs (henceforth  $\text{cop}_1$  and  $\text{cop}_2$ ) occupy the same clause, which should not be possible under standard theories of clausal architecture.

- (1) The issue **is, is** we're out of money.

Double-*is* occurs frequently in spoken discourse and is gaining ground as a conventionalized written idiom (Andersen 2002), attracting the attention of linguists, grammar mavens, and casual observers alike. It occurs across registers and national and socio-demographic speech communities. For example, one of a handful of Language Log posts observing double-*is* behavior documents President Obama's frequent use of double-*is* over the course of three Presidential debates with Governor Romney.<sup>1</sup> Despite its high visibility, there remain relatively few contemporary syntactic treatments of double-*is* (McConvell 2004, Massam 1999, 2013), none of which successfully accounts for all of its properties. The aim of the present paper, therefore, is to provide such an analysis, capable of capturing the relationship between double-*is* and similar specificational copular sentences types.

I propose that *true* double-*is* sentences are Topic-Comment structures (see also Brenier and Michaelis 2005) derived in core syntax.  $\text{cop}_1$  is a topic-marker, and  $\text{cop}_2$  heads FinP. The following illustrates the proposed structure of double-*is* sentences:

- (2) [<sub>TopP</sub> The issue<sub>i</sub> [<sub>Top'</sub> is [<sub>FinP</sub> pro<sub>i</sub> [<sub>Fin'</sub> is [<sub>SC</sub>... we're out of money]]]]]

The syntactic constraints on double-*is* sentences can be accounted for by positing that  $\text{cop}_1$  and  $\text{cop}_2$  are base-generated in the left periphery of a clause projected directly from Fin<sup>0</sup>, and by placing the precopular constituent in a dedicated topic position. Double-*is* sentences are part of a family of copular “amalgam” structures, where a finite copula relates a root clause and a predicate over propositions, a (concealed) question, as in (3).<sup>2</sup>

- (3) He's a fool is what he is.

Despite the similarities between double-*is* and copular amalgams, I show that since *is*-doubling occurs in some copular amalgam sentences, a superficially attractive analysis giving the two sentence types identical structures is untenable (contra Massam 1999, 2013): double-*is* sentences must involve *more* structure than amalgams.

The paper is organized as follows. Section 2 gives an overview of the properties of double-*is*. Section 3 lays out the present Topic-Comment approach to double-*is*, building on a structure proposed for copular amalgams. In section 4, I analyze the morpho-syntactic properties of double-*is* sentences using the Topic-Comment model. Next, section 5 describes the results of an acceptability survey which takes a closer look at the structure shared by double-*is* and amalgam pseudoclefts, and briefly discusses some “problem cases”, which require a different treatment. Finally, section 6 concludes the paper.

## 2 Overview of double-*is* sentences

The following gives a brief profile of double-*is* sentences. Data for this paper are drawn from several sources: previous literature, the Corpus of Contemporary American English (COCA, Davies 2008),

\*Thanks to Marcel den Dikken, Christina Tortora, and William Haddican for helpful discussion and advice on this project. Any errors are my own.

<sup>1</sup>Zimmer, Ben. 2012. Obama's 'is is'. Language Log. <http://languagelog.ldc.upenn.edu/nll/?p=4269>.

<sup>2</sup>Although I do not analyze such sentences as actual syntactic amalgamations, I will continue to refer to them with the term “amalgam” for convenience.

informal web searches, and original constructed examples judged by native speakers. Examples for which no attribution is given are original.

## 2.1 Syntactic and interpretive features of double-*is*

Double-*is* sentences include two major constituents, situated on either side of the cop<sub>1</sub>-cop<sub>2</sub> string. The precopular constituent is a proposition-denoting DP or CP. DPs in this structure are typically headed by content nouns denoting propositions, e.g., *thing*, *problem*, *issue*. CPs in the structure are *wh*-clauses whose variable may either correspond directly to the post-copular proposition, as in (4a), or to a subconstituent of the post-copular proposition (4b).<sup>3</sup> The former type of CP may be either an indirect question or a free relative (in many cases, these cannot be disambiguated), while the latter is unambiguously an indirect question (den Dikken et al. 2000).

- (4) *CP-initial double-is*
- a. What<sub>*i*</sub> she said is, is [that she was hungry]<sub>*i*</sub>.
  - b. What<sub>*i*</sub> she wants is, is she wants [a sandwich]<sub>*i*</sub>.

The post-copular expression denotes a proposition, although its overt syntactic form varies. It may be a bare (*that*-less) clause, a clause introduced by *that*, a non-finite clause, a root clause, or a phrase implying a proposition.

- (5)
- a. The problem is, is [I don't like pork].
  - b. The problem is, is [that I don't like pork].
  - c. My hope is, is [to get a vegetarian meal].
  - d. The question is, is [what am I going to eat]?
  - e. The problem is, is [the pork]. (*i.e.*, *some fact associated with the pork*)

Finally, double-*is* sentences contain two occurrences of the finite copula, typically of the form *is* or *was*. In *true double-is*, neither cop<sub>1</sub> nor cop<sub>2</sub> can combine with negation, modals, aspectual auxiliaries, plural agreement, or temporal adverbials:

- (6) What he did {is / was / \*had been / \*could be / \*isn't}, {is / was / \*had been / \*could be / \*isn't} he stole my bag.

These restrictions are examined more closely in section 4, which analyzes cop<sub>1</sub> and cop<sub>2</sub> as the spell-outs of functional heads in the left periphery.

Double-*is* sentences always have a specificational interpretation: the post-copular constituent provides a value for the variable introduced by the precopular constituent (7). A predicative expression cannot occur in the post-copular position (8).

- (7) The issue is, is [that he forgot to order food].
- (8) \*The issue is, is [unfortunate].

Like *wh*-initial specificational pseudoclefts, double-*is* sentences have a fixed topic-focus information structure (9).

- (9) A: Traveling all the time must be the best thing about your job.  
#B: The **worst** thing about my job is, is that I travel all the time!

Although the post-copular constituent in *wh*-initial specificational pseudoclefts and double-*is* sentences is always focal (which follows from predicate inversion, den Dikken 2006), the type of topic

<sup>3</sup>Throughout the paper, I use "DP" to refer to phrases whose surface form is that of a DP, and "CP" for phrases that are introduced by a complementizer or *wh*-element, whether they are free relatives (formally DPs) or indirect questions (formally CPs).

interpretation of the precopular constituent differs. In pseudoclefts, the *wh*-clause indicates a pre-supposed open question, while the precopular constituent in a double-*is* sentence can establish a new discourse referent (Brenier and Michaelis 2005). Since *cop*<sub>2</sub> in both pseudoclefts and double-*is* sentences serves to introduce a focus, which it may also do in ordinary predicational contexts, the interpretive difference between the sentence types points to *cop*<sub>1</sub>—the element that distinguishes them—rather than *cop*<sub>2</sub>, as the signal of special information structural content (contra, e.g., Massam 1999, Coppock and Staum-Casasanto 2004).

Double-*is* sentences are subject to stringent syntactic restrictions, in addition to the morphosyntactic restrictions on the form of *cop*<sub>1</sub> and *cop*<sub>2</sub>. Their syntax is “frozen”, in the sense that they resist A- and A'-extraction. In addition, double-*is* clauses have an extremely limited distribution in embedded contexts, occurring only under bridge verbs (see section 4).

## 2.2 Double-*is* and pseudoclefts

Given the similarities in form and interpretation between double-*is* and specificational pseudoclefts, it is tempting to analyze double-*is* as a reduced specificational pseudocleft, e.g., with a null or deleted *what* or with the precopular noun itself binding an empty complement to *cop*<sub>1</sub> in a “set-up clause” in the subject position of *cop*<sub>2</sub> (Massam 1999). There are a number of differences between double-*is* and pseudoclefts, however (see, e.g., Coppock and Staum-Casasanto 2004, Massam 2013).

Perhaps the most salient difference between double-*is* sentences and specificational pseudoclefts is that the latter are reversible (10), while the former are not (11).

- (10) a. What he ate is an apple.  
       b. An apple is what he ate.
- (11) a. The problem is is (that) he left early.  
       b. (That) he left early is ⟨\*is⟩ the problem ⟨\*is⟩.

The fixed order of double-*is* sentences is expected if they are Topic-Comment structures, as proposed in section 3.

In addition, double-*is* most frequently features the DP *the thing* in its precopular position, which does not occur in a *what*-ful specificational pseudocleft (Brenier and Michaelis 2005).

- (12) a. ??What the thing is is that there's nothing else to buy. (Brenier and Michaelis 2005:53  
       (15))  
       b. The thing is, is that there's nothing else to buy.

Finally, double-*is* in fact occurs robustly in *wh*-initial pseudoclefts (13a). If double-*is* were a *what*-less pseudocleft, the source for double-*is* pseudoclefts would have to feature a nested free relative in subject position, which is unlikely, given the fact that double-*is* pseudoclefts like (13a) impose no particularly heavy processing load, while nested free relatives like (13b) do.

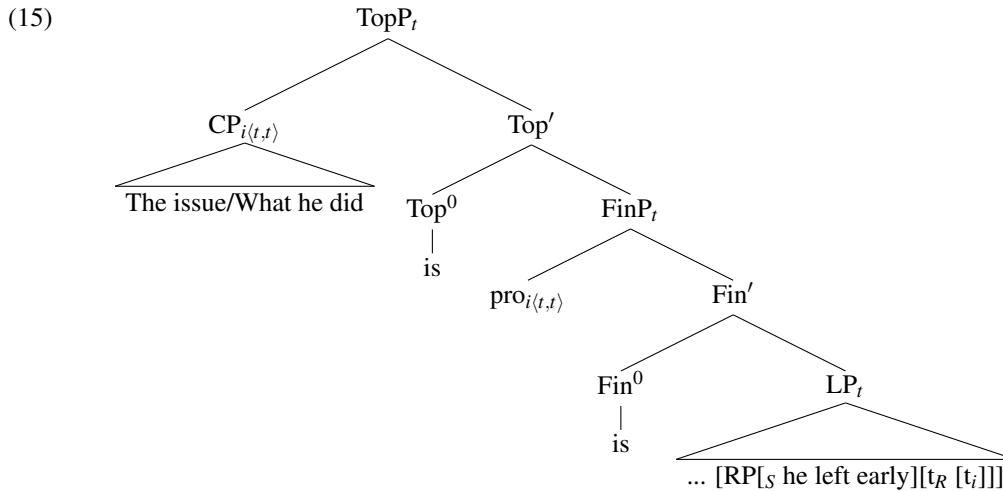
- (13) a. What I want is, is I want a sandwich.  
       b. ?What what I want is, is I want a sandwich.

The precopular constituent of double-*is* can be an indirect question (e.g., with multiple *wh*-expressions) (14a). While (13b) is not ill-formed, (14b) is. Embedding an indirect question in a free relative or another indirect question is not possible, and so a reduced pseudocleft is not the source for (14a).

- (14) a. Who ate what is, is John ate a sandwich and Mary ate a salad.  
       b. \*[What [who ate what] is], is John ate a sandwich and Mary ate a salad.

### 3 A Topic-Comment proposal for *true double-is*

I propose that double-*is* sentences are Topic-Comment structures (cf. den Dikken et al. 2000 on amalgam specificational pseudoclefts), where  $\text{cop}_1$  heads  $\text{Top}^0$  and  $\text{cop}_2$  heads  $\text{Fin}^0$ :



The topic phrase licenses a coreferential *pro*-predicate in  $\text{Spec,FinP}$ , which is predicated of the focused post-copular clause.<sup>4</sup>

Double-*is* sentences instantiate a type of clause in English, the copular amalgam, which is headed by a finite copula ( $\text{cop}_2$ ) base-generated in a CP domain that does not dominate a lower TP or VP (O’Neill 2012, in prep.). In addition to a “main” copula in  $\text{Fin}^0$ , double-*is* sentences feature an additional copula in  $\text{Top}^0$  which explicitly marks the precopular constituent as a topic; thus, information structure in double-*is* sentences is fixed.

The English copula *be* has the unique freedom to distribute in any [+V] syntactic environment, including the clausal left periphery, because of its lack of lexical semantic content. It serves simply as a vacuous predication functor (Partee 1986) combining a lower-type element  $X$  with a predicate of type  $\langle X, t \rangle$ . Given this polymorphism, its argument  $X$  can be a proposition (type  $t$ ), such that  $\langle X, t \rangle$  is a predicate over propositions (abstracting away from intensionality here). As in other specificational copular sentences, which carry an exhaustivity implicature, the predicate is a singleton set. Since the logical subject, the unique member of the set, is a focused proposition, the predicate has the form of a(n indirect or concealed) question (see also Schlenker 2003, Caponigro and Davidson 2011 on question-answer clauses).

The copula can spell out the morphological features of any functional head compatible with its predicational semantics, provided no lexical verb is available to do so and all expressions in the sentence can be independently licensed. Section 2 noted that the copulas in double-*is* cannot combine with material in the T/Infl domain of the clause (henceforth “baggage”) other than simple  $[\pm\text{past}]$ ; I take this as evidence that the copulas of double-*is* do not occupy T (or V+T), unlike the copula of standard copular sentences. I propose a finite clause in English can be projected directly from a functional head in the C domain that is *not* associated with lower T or V-projections. I assume that when this happens, the copula ( $\text{cop}_2$  in double-*is*) spells out  $\text{Fin}^0$ , the lowest head in the complementizer domain, in order to realize the feature [+finite] and stranded [tense] features (Bjorkman 2011, O’Neill in prep.). These Fin-headed clauses lack the structure necessary for nominal argument licensing (i.e., a phi probe on T), so their (logical) subject argument is always non-nominal: it is an embedded root clause.

Double-*is* sentences have been compared to copular amalgam sentence types before. While I do not treat copular amalgams/Fin-headed clauses as actual syntactic amalgamations, I analyze all

<sup>4</sup>The *pro*-predicate ends up in  $\text{SpecFinP}$  as a result of predicate inversion mediated by a Linker Phrase, in the sense of (den Dikken 2006). The focus clause is a small clause subject. I have left the representation of predicate inversion in the tree structure vague for reasons of space.

double-*is* sentences as containing the same base structure as the Fin-headed sentence types illustrated in (16)–(19) (for more on these sentence types, see, e.g., McConvell 1988, 2004, Massam 1999, 2013, den Dikken et al. 2000, Brenier and Michaelis 2005, Ross-Hagebaum 2004, Lambrecht and Ross-Hagebaum 2006, O’Neill 2012, in prep.):

- (16) Hypotactic amalgam
  - a. That’s the main thing is that you can’t tell. (Brenier and Michaelis 2005:76 (45a))
  - b. That’s what I was about to say is that everyone needs to be tested. (Ross-Hagebaum 2004:403 (4))
- (17) Amalgam pseudoclefts
  - a. What we need now is we need more data.
  - b. What he did is he called her back.
- (18) Reverse amalgam pseudoclefts
  - a. We need more data is what we need.
  - b. He called her back is what he did.
- (19) Anchored intrusive *be*
  - a. I want to recommend something that might help is that you must say to them when you are available before the committee is struck. (Massam 2013:6 (18b))
  - b. Okay, so maybe I should write about that in the first paragraph, is the fact that the biological parent is the only one who has standing. (O’Neill 2012:35 (82))

Crucial for the present proposal is the fact that all of the preceding sentence types support double-*is*, save the reverse amalgam pseudocleft (18), which is the only one where the focus rather than the topic occurs in precopular position. Note, for example, (20):

- (20) ...that’s much the problem with reporting on Pakistan **is is** those short news items... (COCA)

#### 4 Syntactic analysis of double-*is*

The Topic-Comment analysis of double-*is* sentences makes a number of predictions, which are borne out in the data reported in previous literature, and corroborated by findings from an original search of 85 million words of spoken English from 1990–2010 in the Corpus of Contemporary English (COCA, Davies 2008). I searched COCA for all and only sentence tokens with two adjacent occurrences of the copula, where one of them appeared syntactically extraneous, and found 514 double-*is* tokens. The search included tokens in which *cop*<sub>1</sub> or *cop*<sub>2</sub> was associated with baggage.

##### 4.1 Deriving the restrictions on double-*is* sentences

**Embedding** Double-*is* sentences can only be embedded under bridge verbs (21), which allow embedded topics. The relative unembeddability of double-*is* follows under the present analysis: they have “too much” structure (i.e., TopP) in their left periphery to be embedded in non-bridge environments (22).

- (21) a. We decided, by the way, that the only way they’re going to let the news media count the ballots in the state of Florida was is if they threw you and Bob in the same room together and let you do it and that would be acceptable. (COCA)
- b. *Compare*: We decided that beer, we would never drink.
- c. I guess the question is, is is it right to use our open seas as a testing lab? (COCA)
- d. *Compare*: I guess pizza, we could get at the store.
- (22) a. \*She regrets that her mistake is, is that she forgot her lunch.
- b. \*I consider my mistake { $\emptyset$  / to be}, { $\emptyset$  / to be} that I forgot my lunch.

**Extraction** Double-*is* sentences are syntactically “frozen”; for example, the precopular constituent cannot be questioned (23a). In this respect, they are similar to specificational pseudoclefts (23b).

- (23) a. \*What<sub>*i*</sub> do you think t<sub>*i*</sub> is, is we're hungry? —The problem.  
 b. \*What<sub>*i*</sub> do you think is an apple? —What she ate.

Extraction of the post-copular constituent is impossible, as is subextraction from both the precopular and post-copular constituents (25).

- (24) a. \*What<sub>*i*</sub> do you think the problem is, is t<sub>*i*</sub>? —That he forgot his lunch.  
 b. \*What<sub>*i*</sub> do you think that what she ate is t<sub>*i*</sub>? —An apple.  
 (25) a. \*[How big]<sub>*i*</sub> do you think [the t<sub>*i*</sub> problem] is, is that he forgot his lunch?  
 b. \*What<sub>*i*</sub> do you think the problem is, is that he forgot t<sub>*i*</sub>?

Frozensness in both pseudoclefts and double-*is* can be attributed to information structure (see discussion in den Dikken 2006: Ch. 4.4). The post-copular expression is already the focus as a result of predicate inversion; since the precopular expression is a topic, it cannot be assigned a focus interpretation in an A' environment, nor is it possible to subextract out of the focal constituent.

The amalgam pseudocleft example in (26) below and the double-*is* sentence in (25b) above are strikingly worse than the others, because they are syntactically, and not just information-structurally ill-formed.

- (26) \*What<sub>*i*</sub> do you think that what she ate is she ate t<sub>*i*</sub>? —An apple.

The analysis in the present paper takes them to involve subextraction out of the same left branch position: the propositional subject of the amalgam-type copular sentence, which occupies the underlying subject position of a small clause.

#### 4.2 Deriving the restrictions on the copulas

Central to the present proposal is the claim that although cop<sub>2</sub> is *finite*, it cannot combine with other material from the V or T domain, because it heads a FinP clause which does not dominate a TP or VP. Like previous accounts of double-*is* treating cop<sub>2</sub> as the spell-out of a left peripheral functional head (Massam 1999, 2013, Coppock and Staum-Casasanto 2004), I predict that cop<sub>2</sub> can only take the simple finite form *is* or *was*. Cop<sub>1</sub> is predicted to behave similarly, because it is simply the spell-out of a topic-marker. The former prediction is borne out: baggage on cop<sub>2</sub> is rejected in acceptability judgments and almost unattested in COCA; however, cop<sub>1</sub> does show some flexibility.<sup>5</sup> A closer look at the relationship between double-*is* and amalgams (see section 5) reveals

	Bare	<i>being</i>	Modal + <i>be</i>	Aux + <i>been</i>	<i>are</i>	Other
Cop <sub>1</sub> :	454	29	15	6	5	5
Cop <sub>2</sub> :	512	0	0	1	1	0

Table 1: Form of cop<sub>1</sub> and cop<sub>2</sub> in COCA

that sentences containing a cop<sub>1</sub>-cop<sub>2</sub> string can arise from two underliers.

The most robust restriction on the copulas is that neither cop<sub>1</sub> nor cop<sub>2</sub> can be negated:

- (27) {The issue is(\*n't), is(\*n't)} / {What bothers me is(\*n't), is(\*n't)} that I'm hungry.

Cop<sub>2</sub> cannot support modals, auxiliaries, or plural agreement:

- (28) a. \*The issue is, could be he left his bag at home.  
 b. \*What she thought was, had been that you had done it already.  
 c. \*What he sees is/are, are he sees disasters and problems everywhere.

<sup>5</sup>There is one token of cop<sub>2</sub> as *are*, following an occurrence cop<sub>1</sub> as *are*. This lone token may well be a disfluency.

Modals, auxiliaries, and plural agreement may, however, occur on  $\text{cop}_1$ , which is unexpected if it is a simple topic-marker (see Table 1), but crucially, such baggage is only possible if the precopular constituent is a DP. I will return to this observation shortly.

- (29) a. The issue could be, is he forgot his bag.
- b. \*What he forgot might be, is he forgot his bag.
- (30) a. The cruel facts of life are, is that not every person who teaches Art is a good artist himself. (Coppock and Staum-Casasanto 2004:2 (9a))
- b. \*What she counts on are, is that you'll get there on time and that you'll do your job.

### 4.3 Summary

Generally, the syntactic features of double-*is* are predicted by the Topic-Comment analysis, where  $\text{cop}_1$  is a topic-marker and  $\text{cop}_2$  heads  $\text{Fin}^0$ . As expected, double-*is* sentences are unembeddable in non-bridge contexts, because they have a rich left periphery. The two major constituents always occur in the variable-value order, which is consistent with their fixed information structure. None of the examples in COCA featured extraction of or out of the two major constituents of double-*is*, because of both information structure and left-branch extraction restrictions. Since  $\text{cop}_2$  is a simple  $\text{Fin}^0$  head, the form of  $\text{cop}_2$  must be morphologically bare.

## 5 Two sources for double-*is*: a closer look double-*is* in amalgams

The fact that  $\text{cop}_1$  shows some variation in morpho-syntactic form requires closer analysis. At first glance, it seems to contradict the analysis demonstrated in (15), where  $\text{cop}_1$  is a simple topic-marker.

While  $\text{cop}_2$  is nearly always bare, in 60 tokens,  $\text{cop}_1$  had a non-bare form. 29 of these non-bare tokens involved the non-finite form *being* (e.g., *The thing being is that...*). Since this work is concerned with the unusual distribution of the *finite* copula, *thing being* sentences will not be analyzed further. Interestingly, the form of  $\text{cop}_1$  only combines with baggage when the precopular constituent is a plain DP, and never when it is a CP, as in a double-*is* pseudocleft. To my knowledge, this correlation has not been noticed before.

	Total	$\text{Cop}_1$ bare	$\text{Cop}_1$ <i>being</i>	$\text{Cop}_1$ baggage
Total	514	454	29	31
DP double- <i>is</i>	452	392	29	31
CP double- <i>is</i>	62	62	0	0

Table 2: Precopular constituent and form of  $\text{cop}_1$  in COCA.

The interaction between the category of precopular constituent and the morphological form of  $\text{cop}_1$  provides key insight into the structure of *true* double-*is* sentences, which are built from  $\text{Fin}$ -headed clauses. Double-*is* can co-occur with amalgam copular sentences; in this structure, the precopular constituent, a complete sentence or a CP, occupies a topic position (31).

- (31) a. What I need i  
s, is I need a cup of coffee.
- b. [That's what I need] is, is I need a cup of coffee.

When  $\text{cop}_1$  occurs in an amalgam, the present analysis predicts that it must take a bare form, since it occupies a simple functional head:  $\text{Top}^0$ . Co-occurrence of double-*is* with an amalgam does not make available any ordinary verbal position for  $\text{cop}_1$ . Any verbal element showing baggage in this environment must be located inside the topicalized clause preceding  $\text{cop}_1$ . While the fact that baggage on  $\text{cop}_1$  is totally unattested in COCA with CP-initial double-*is* is suggestive, the relatively low frequency of CP-initial double-*is* makes the evidence imperfect. This section presents an acceptability survey showing that baggage on  $\text{cop}_1$  double-*is* amalgam pseudoclefts not merely unattested, but actually unacceptable.

A survey was constructed to compare the acceptability of baggage on  $\text{cop}_1$  in double-*is* amalgam pseudoclefts to that of baggage on the sole copula of single-*is* amalgam pseudoclefts. Assuming that the Topic-Comment structure illustrated in (15) and the FinP it contains are accurate, neither copula in a true double-*is* sentence nor the copula in the amalgam pseudocleft should be able to combine with baggage. The experiment consisted of a self-paced acceptability judgment survey hosted by Ibex farm, completed by 20 native speakers of English from the United States (median age=31), recruited via Amazon Mechanical Turk. Participants were instructed to rate the “naturalness” of each sentence if it were used in a casual conversation, using a 9-point Likert scale. The survey items were divided into four conditions (see Table 3). Four versions of the survey were constructed, with 16 experimental items and 16 filler counterbalanced by Latin square.

	Single <i>is</i>	Double <i>is</i>
Bare	What he likes is he likes pizza	What he likes is, is he likes pizza
Baggage	What he likes could be he likes pizza	What he likes could be, is he likes pizza

Table 3: Factor design for acceptability judgment survey

A linear mixed effects regression model with random intercepts for item and subject was used to test for a main effect of baggage. An ANOVA between the *lmer* model and the null model showed a main effect of baggage overall ( $p < 0.0001$ ) and within each sentence type. As expected, there was no interaction between the factors, because baggage has a similar effect on  $\text{cop}_1$  ( $p = 0.0009$ ) and  $\text{cop}_2$  ( $p = 0.0028$ ); that is, it lowers acceptability in both single-*is* amalgam specificational pseudoclefts (with  $\text{cop}_2$ ) and double-*is* amalgam specificational pseudoclefts (with baggage on  $\text{cop}_1$ ).<sup>6</sup>

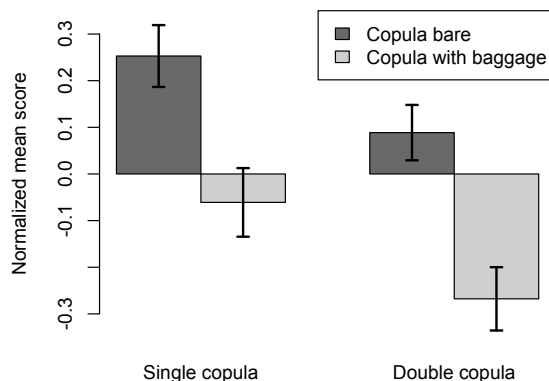


Figure 1: Results of acceptability survey

The results of this experiment indicate that a string containing multiple occurrences of the copula can come about in more than one way. In *true* double-*is*, which is the only structure available to double-*is* copular amalgam sentences,  $\text{cop}_1$  is a topic-marker and  $\text{cop}_2$  heads  $\text{Fin}^0$ . When  $\text{cop}_1$  combines with baggage, however, it cannot be a simple topic-marker, so sentences such as (32) need a different analysis.

(32) [The only difference could be] is that you want to give more state control... (COCA)

I propose that such sentences are simple Fin-headed clauses, with a copular clause occupying SpecFinP (the “set-up clause”). Their  $\text{cop}_1$  is not a topic-marker, but the main verb of the set-up clause. As such, it can combine with a fully articulated T domain. Like the verbal copula of the specificational pseudocleft,  $\text{cop}_2$  always relates a presupposed constituent harboring a variable to a focal constituent supplying the value for that variable, so the set-up clause in (32) is itself a specificational copular sentence, where the value—the underlying subject—is a null element.

<sup>6</sup>Double-*is* sentences had lower acceptability overall than non-doubled sentences, but the normalized mean was positive. I attribute the lower ratings to the garden-path effect of reading three adjacent finite verb forms (the lexical verb of the pseudocleft, followed by two occurrences of the copula), and to possible prescriptive bias.



So far, this proposal is very much in line with the structure Massam (2013) develops for *all* double-*is* sentences and amalgams. She treats cop<sub>2</sub> as an Appositive head (Appo<sup>0</sup>), an applicative element that takes an “anchor” in its specifier (a null category bound by the precopular DP or CP) and the focal post-copular constituent (which specifies the content of the anchor) in its complement.

- (33) [<sub>IP</sub> [<sub>DP<sub>i</sub></sub> the thing] [<sub>I'</sub> [<sub>I</sub>] [<sub>VP</sub> [<sub>V'</sub> [<sub>V</sub> is ] [<sub>AppoP</sub> [ e<sub>i</sub> ] [<sub>Appo'</sub> [<sub>Appo</sub> is ] [<sub>CP</sub> that I like you]]]]]]]]

As we have seen, this structure cannot account for *true* double-*is*, since cop<sub>2</sub> must occupy Appo<sup>0</sup>, leaving no room for cop<sub>1</sub> when double-*is* occurs with an amalgam sentence type. Massam’s (2013) applicative appositive structure admits a position for one “intrusive” copula, but not two, so the Topic-Comment model of *true* double-*is* developed in the present paper is needed.

The present Fin-headed clause proposal, unlike the apposition model, analyzes cop<sub>2</sub> as relating full propositions. Since the anchor in the set-up clause need not be adjacent to the copula, this analysis can straightforwardly account for examples like (19b), in section 3, where they are non-adjacent. I propose that the anchor in these putative double-*is* sentences with baggage on cop<sub>1</sub> is a null operator occupying a position above the fronted predicate, e.g., *the thing* in (33).<sup>7</sup> The proposed structure for the set-up clause in sentences like (32) and (33) is given in (34).

- (34) [<sub>CP</sub> Op<sub>i</sub> [<sub>C'</sub> [<sub>C</sub>] [<sub>TP</sub> [<sub>DP<sub>i</sub></sub> the thing] [<sub>T'</sub> [<sub>T+V</sub> is] [<sub>VP</sub> t<sub>V</sub> [<sub>LP</sub> t<sub>k</sub> [<sub>L'</sub> [<sub>L+R</sub>] [<sub>RP</sub> e<sub>i</sub> [<sub>R'</sub> [<sub>t<sub>R</sub>] [<sub>DP</sub> t<sub>k</sub>] ... ]]]]]]]]]]]</sub>

In most contexts, null operator movement over a topic is barred, as it would yield an information structure clash, but it is not necessarily ill-formed. Extraction of an *overt* operator across a fronted predicate brings about a clash, since the operator corresponds to a focus in its base position (as the subject of specification), but is associated with a topic interpretation in its landing site. A *null* operator receives no local information structural interpretation, so it avoids this problem. The null operator must, however, receive content from somewhere. If the constituent providing its content is a topic, once again, a clash arises, since the null operator is syntactically associated with the focus of specification:

- (35) a. #The burgers<sub>i</sub> Op<sub>i</sub> that [what he ate] is e<sub>i</sub> were greasy.  
 b. #Those burgers<sub>i</sub> are hard to believe Op<sub>i</sub> that [what he ate] is e<sub>i</sub>.

In the set-up clause of an amalgam, by contrast, the operator is identified only with the content of the post-copular constituent, which is a focus. The operator therefore receives one coherent information structural interpretation in the derivation: it is the focus of its local specificational clause, and its post-copular associate is the focus of the larger specificational structure. This recursive specificational structure is more complex than the *true* double-*is* structure, where the precopular constituent is a base-generated topic co-indexed with a fronted predicate. I speculate that, although all double-*is* sentence strings with simple DPs or free relatives in precopular position are ambiguous between the recursive specificational structure and the more direct Topic-Comment structure, the latter is preferred, and thus baggage occurs infrequently.

A further prediction of the present proposal deriving double-*is* strings from two difference structures is that the *true* double-*is* structure should be able to contain the other, resulting in a string of *three* copulas. Indeed, three copular forms can occur in a row (36). Such examples have been noted in the previous literature, but not incorporated into a unified analysis.

- (36) a. The fact is is Howie is that with a quote like that, you press the subject...<sup>8</sup>  
 b. [The problem remains] is<sub>cop1</sub>, is<sub>cop2</sub> that I haven’t eaten.  
 c. ...[One reason **could be**] is<sub>cop1</sub>, is<sub>cop2</sub> that I’m as addicted to him as he is to his DOC.  
 (web example)

To summarize this section, both cop<sub>1</sub> and cop<sub>2</sub> in true double-*is* are simple functional heads. Sentences in which cop<sub>1</sub> occurs with baggage actually have a different structure, in which cop<sub>1</sub> is a verb. They share their basic structure with copular amalgams, where cop<sub>2</sub> occupies Fin<sup>0</sup>, and no special topic-marker is projected.

<sup>7</sup>Thanks to Marcel den Dikken for suggesting this line of analysis.

<sup>8</sup>Zimmer, Ben. 2011. The elusive triple “is”. *Language Log*. <http://languagelog.ldc.upenn.edu/nll/?p=3452>.

## 6 Conclusion

This paper has offered a syntactic analysis of the puzzling double-*is* construction in English, relating it to other amalgam copular sentence types, where  $\text{cop}_2$  projects directly from  $\text{Fin}^0$ . In *true* double-*is* sentences,  $\text{cop}_1$  functions as a topic-marker. Some evidence for the status of  $\text{cop}_1$  as  $\text{Top}^0$  is the fact that it cannot occur in sentences where the focus is in precopular position, the fact that double-*is* sentences are unembeddable, and the fact that the major constituents of double-*is* sentences are not reversible. When  $\text{cop}_1$  is not bare, a putative double-*is* sentence is analyzed as a simple amalgam, with a specificational copular set-up clause in SpecFinP. By teasing apart these two sources for double-*is* strings, the present paper accounts for the co-occurrence of double-*is* with other copular amalgams. It also provides a direction for a contemporary syntactic treatment of the understudied copular amalgam sentence type (see O'Neill in prep. for a detailed analysis).

## References

- Andersen, Gisle. 2002. Corpora and the double copula. *Language and Computers* 40:43–58.
- Bjorkman, Bronwyn. 2011. BE-ing default: The morphosyntax of auxiliaries. Doctoral dissertation, Massachusetts Institute of Technology.
- Brenier, Jason M, and Laura A Michaelis. 2005. Optimization via syntactic amalgam: Syntax-prosody mismatch and copula doubling. *Corpus Linguistics and Linguistic Theory* 1:45–88.
- Caponigro, Ivano, and Kathryn Davidson. 2011. Ask, and tell as well: Clausal question-answer pairs in ASL. *Natural Language Semantics* 19:323–371.
- Coppock, Elizabeth, and Laura Staum-Casasanto. 2004. Origin of the English double-is construction. Manuscript, Stanford University.
- Davies, Mark. 2008. The Corpus of Contemporary American English (COCA): 400+ million words, 1990–present. Online: <http://corpus.byu.edu/coca>.
- den Dikken, Marcel. 2006. *Relators and linkers: the syntax of predication, predicate inversion, and copulas*. Cambridge, Mass: MIT Press.
- den Dikken, Marcel, André Meinunger, and Chris Wilder. 2000. Pseudoclefts and ellipsis. *Studia Linguistica* 54:41–89.
- Lambrecht, Knud, and Sebastian Ross-Hagebaum. 2006. Apo koinou and intrusion: Toward a syntactic-pragmatic typology of amalgam constructions in spoken English. In *Proceedings of the 4th workshop on discourse structure*. University of Texas, Austin.
- Massam, Diane. 1999. *Thing is* constructions: the thing is, is what's the right analysis? *English Language and Linguistics* 3:335–352.
- Massam, Diane. 2013. Intrusive *be* constructions in (spoken) English: Apposition and beyond. In *Proceedings of the 2012 annual conference of the Canadian Linguistic Association*.
- McConvell, Patrick. 1988. To be or double be? Current changes in the English copula. *Australian Journal of Linguistics* 8:287–305.
- McConvell, Patrick. 2004. Catastrophic change in current English: Emergent double-be's and free-be's. Talk given at the CRLC Seminar at the Australian National University.
- O'Neill, Teresa. 2012. Coordination in English copular amalgams. Qualifying Paper, CUNY Graduate Center.
- O'Neill, Teresa. In prep. The domain of finiteness: anchoring without Tense in English amalgam copular sentences. Dissertation, CUNY Graduate Center.
- Partee, Barbara. 1986. Ambiguous pseudoclefts with unambiguous *be*. In *Proceedings of the North East Linguistic Society*, volume 16, 354–366.
- Ross-Hagebaum, Sebastian. 2004. The *that's X is Y* construction as an information-structure amalgam. In *Paper presented at the 30th meeting of the Berkley Linguistics Society*, 403–414.
- Schlenker, Philippe. 2003. Clausal equations (a note on the connectivity problem). *Natural Language & Linguistic Theory* 21:157–214.

Linguistics Program  
 The Graduate Center, City University of New York  
 365 5th Avenue, Room 7407  
 New York, NY 10016  
[toneill@gc.cuny.edu](mailto:toneill@gc.cuny.edu)